# Metrics for Quantifying Predictability Limits


# Kathy Pegion

*University of Colorado/CIRES & NOAA/ESRL/PSD*



*Acknowledgements: Matt Newman, Tim DelSole*


*U.S. Clivar PPAI & PSMI Panels July 2014*

# What is the upper limit of skill (i.e. predictability)?

a) Perfect model

b) Cross-model

c) Linear Inverse Model

d) Average Predictability Time

**Perfect Model -- How well does model predict itself?**

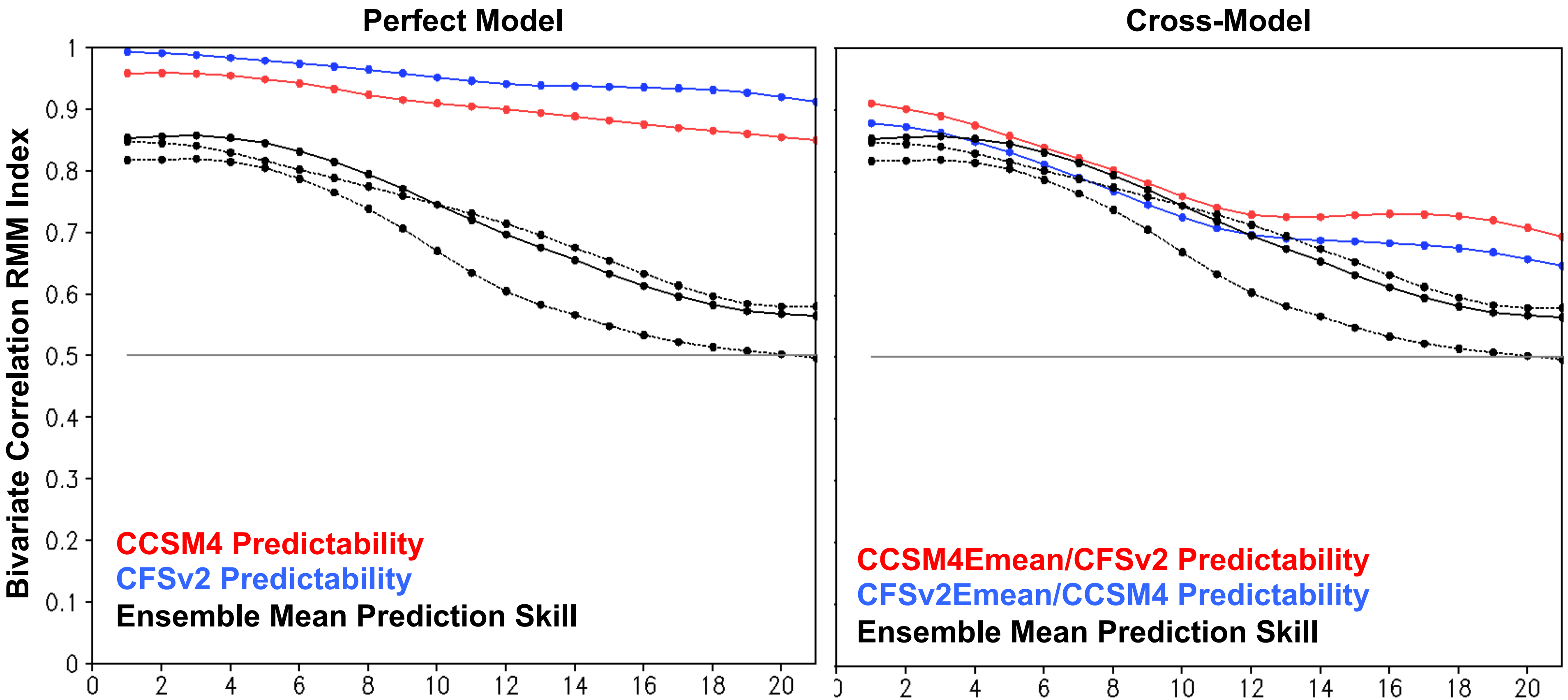
- Withhold one ensemble member and use the ensemble mean of the rest to predict that one member. Repeat for all members.

- Models are very good at predicting themselves (i.e. model spread is too small)

**Cross-Model – How well do models predict each other?**

- Same idea, but treat one model's ensemble member as truth and see how well the other model's ensemble mean predicts it.  Repeat for all models and ensemble members.

# Perfect and Cross-Model Predictability




Pegion, NMME Project

# Using LIM to estimate predictability

$$\mathbf{d\,x / dt} = \mathbf{L\,x} + \mathbf{F_s}$$

$\mathbf{L}$ = constant, $\mathbf{F_s}$ = additive (state-independent) noise.

$$\mathbf{x}(t+\tau) = \exp(\mathbf{L}\tau)\,\mathbf{x}(t) + \varepsilon = \mathbf{G}(\tau)\,\mathbf{x}(t) + \varepsilon$$

**"signal"**

**"noise"**

**Expected forecast error covariance**

(assuming no initial error) :

$$\mathbf{E}(\tau) = < \varepsilon\,\varepsilon^{T} > = \mathbf{C}(0) - \mathbf{G}\,\mathbf{C}(0)\,\mathbf{G}^{T}$$

**Expected forecast anomaly correlation**

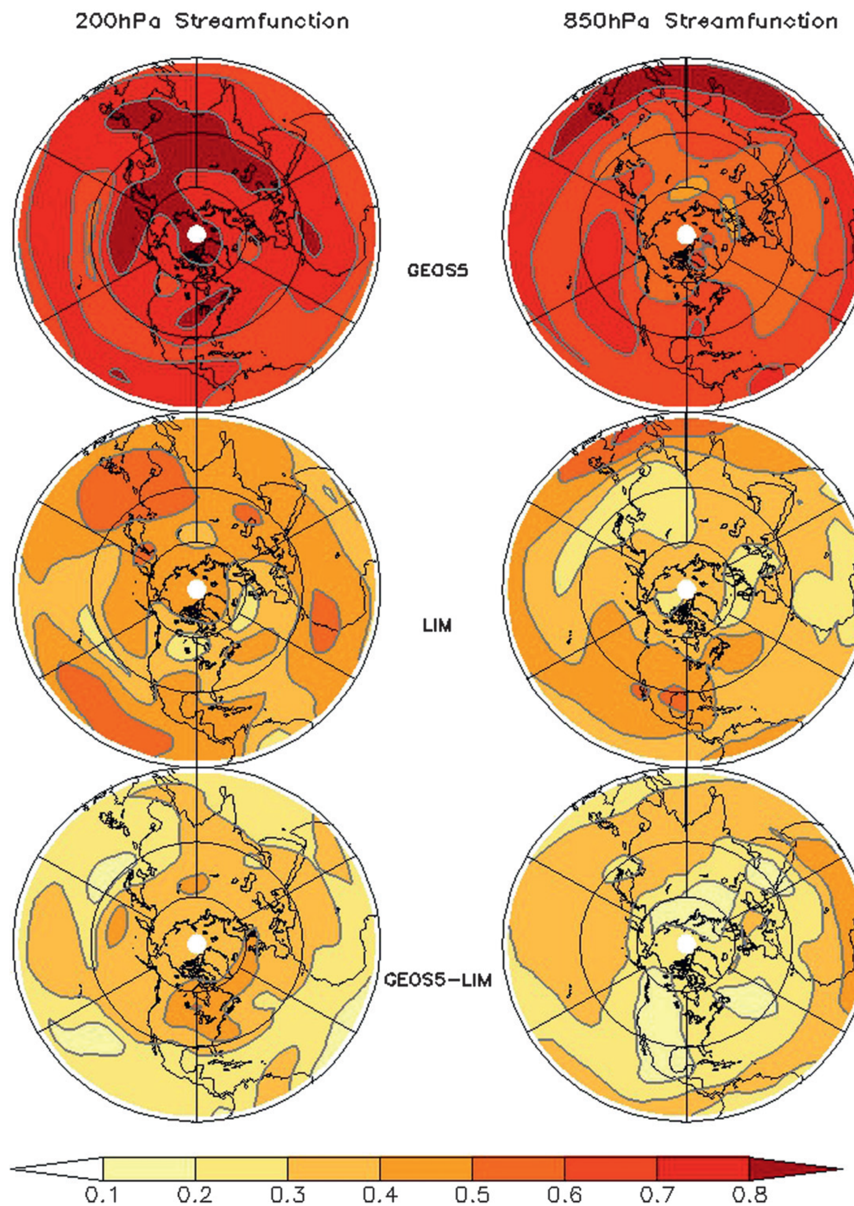
$$\rho_{\infty} = \frac{s}{\sqrt{1+s^2}}, \text{ where } s^2 = \frac{[\mathbf{G}\,\mathbf{C}(0)\,\mathbf{G}^{T}]_{ii}}{[\mathbf{E}(\tau)]_{ii}}$$

**Larger signal related to leading singular vector of G(τ)**

*Courtesy of Matt Newman*

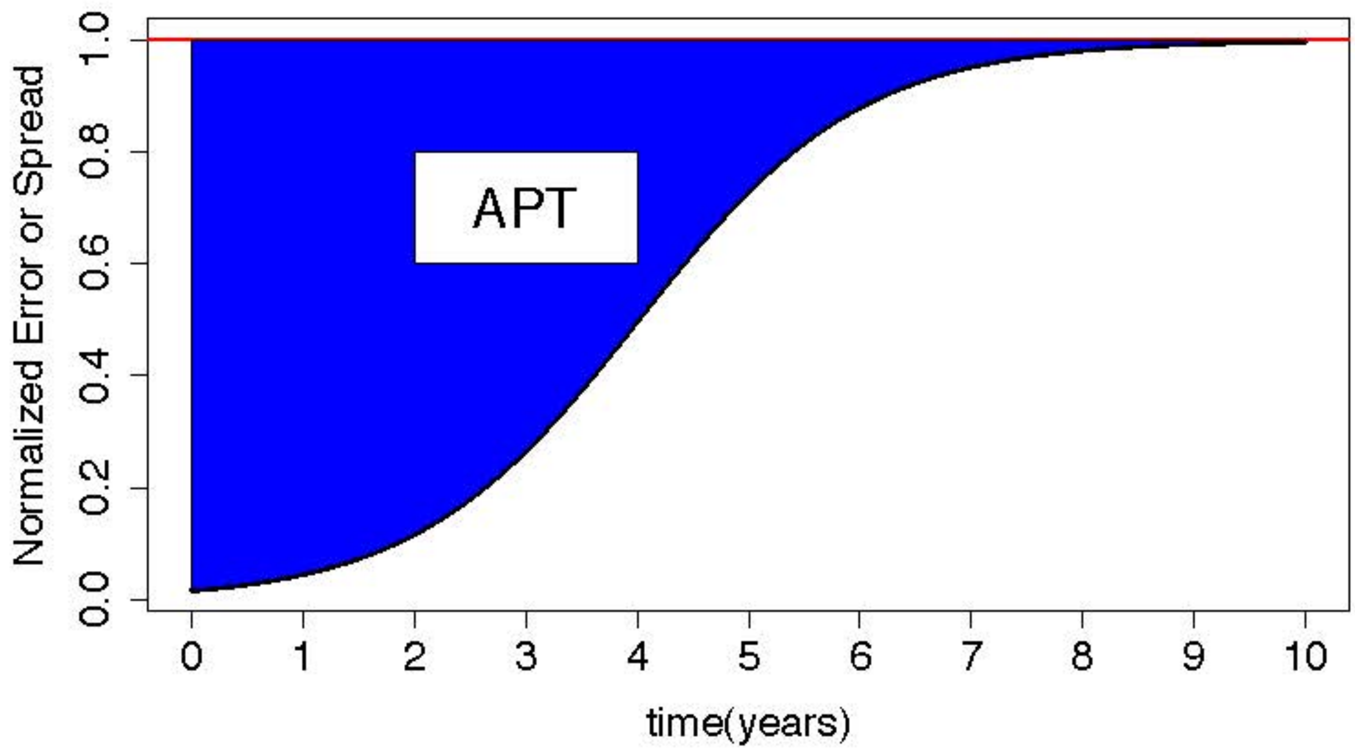
# Linear Inverse Model (LIM)

Week-3
Forecasts




*Pegion and Sardeshmukh 2011, MWR*

# Average Predictability Time




$$APT = 2 \int_0^\infty \left( \frac{\sigma^2_{clim} - \sigma^2_{forecast}(\tau)}{\sigma^2_{clim}} \right) d\tau$$

DelSole and Tippett, 2009a,b, *J. Atmos. Sci.*

*Courtesy of Tim DelSole*

# Optimize APT

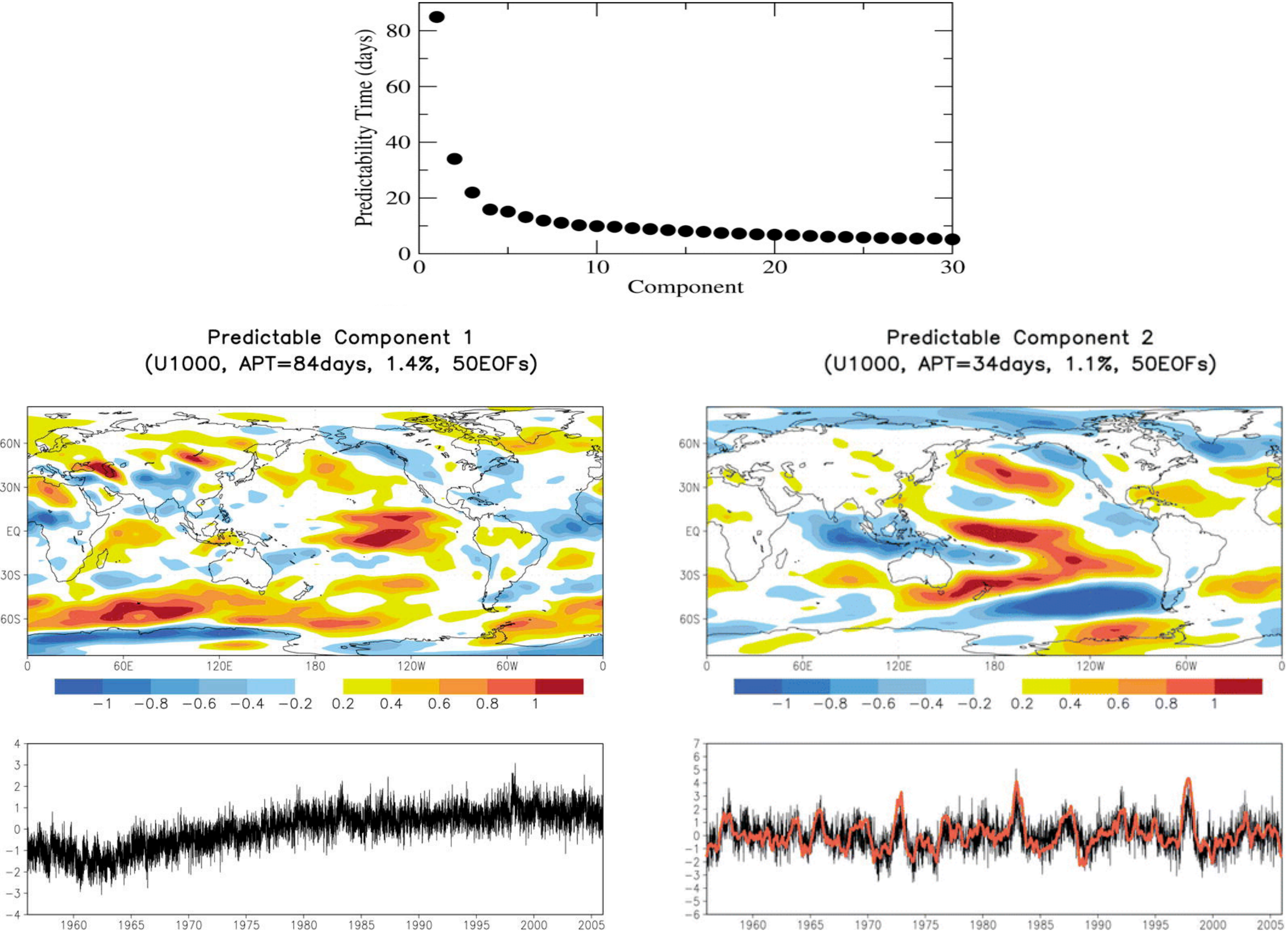
Find the linear combination of data that maximizes APT. Solution:

$$\left(2\int_0^\infty \boldsymbol{\Sigma}_c - \boldsymbol{\Sigma}_f(\tau)d\tau\right)\mathbf{q} = \lambda\boldsymbol{\Sigma}_c\mathbf{q}$$

where $\boldsymbol{\Sigma}_f$ and $\boldsymbol{\Sigma}_c$ are the forecast and climatological covariance matrices.

- Eigenvalue $\lambda$ gives the APT.
- Eigenvectors $\mathbf{q}$ are projection vectors for generating time series.
- Resulting time series are uncorrelated in time.
- Each projection vector is associated with physical pattern $\mathbf{p} = \boldsymbol{\Sigma}_c\mathbf{q}$.
- Physical pattern $\mathbf{p}$ is called a predictable component.
- Product of $\mathbf{p}$ * (time series), summed over all components, recovers original time series.

*Courtesy of Tim DelSole*




*from DelSole and Tippett 2009, JAS*

## Noise-Limited

**We are at the predictability limit**

*A forecast of opportunity approach*

- Can we identify *a priori* times of potentially higher forecast skill?
- Do models already pick up on this and is it contained in the ensemble?
- Are we able to exploit this information?

## Science-Limited

**(a) There is skill left to be realized, but we don't know how to do it.**

**(b) We don't know for sure what the limit is**

*A process-based approach*

- Can models represent the relevant processes and phenomena?
- Do we understand why or why not?
- Where is our scientific understanding of these processes and what their predictability limit should be?
- How can we make the leap from finding errors in models to identifying real solutions to fixing those errors?